

COMPLEMENTIZER *PA*, THE FINITENESS OF ITS COMPLEMENTS, AND SOME REMARKS ON EMPTY CATEGORIES IN PAPIAMENTO

Silvia Kouwenberg
University of Amsterdam

The following paper discusses clausal complements in Papiamentu which are introduced by the complementizer *pa*. It will be shown that the finite/nonfinite distinction applies to these and accounts for the distribution of overt subjects and empty subjects in the *pa*-clauses. Furthermore, a discussion of empty category subjects will show that Papiamentu distinguishes between controlled PRO and arbitrary usage of PRO as subjects of infinitival clauses, and allows for the use of small pro subjects in finite clauses in certain highly restricted cases.

1. Introduction

Papiamentu (henceforth P) *pa* has several functions: it is used both as a complementizer and as a preposition, and it is believed that the first function historically derives from the second (cf. Washabaugh 1975). This paper will only be concerned with the function of *pa* as a complementizer.

It is often assumed that creole languages distinguish between complementizers introducing factive and nonfactive complements, the latter being complements of uncertain realization. P *pa*, then, would introduce clausal complements of uncertain realization, in contrast with other complementizers like *ku*. Some creolists assume that *pa* introduces nonfinite clauses only, e.g., Bickerton (1984:182). Others take it that creole languages do not have a finite/nonfinite distinction: Dijkhoff & Mufwene (1986, 1988) adopt this position and attribute the “uncertain realization” reading to the so-called future orientation of *pa* (future orientation is qualified by them as being associated with intent or purposiveness). In the fol-

lowing, it will be shown that neither of these positions represents the P data correctly.

The data is analyzed under the standard theoretical assumptions of Government-Binding theory (GB). We will be referring in particular to the typology of empty categories as outlined in Chomsky (1982), and to Rizzi's (1986) description of pro-drop phenomena, which will be discussed in some detail below. The finite/nonfinite distinction in *pa*-complement clauses will be discussed in section 2. In section 3 we will look in somewhat more detail at the licensing of the empty category subjects ([ec] subjects) of infinitival *pa* clauses. A general discussion of empty categories of the [+ pronominal] type will show that P distinguishes between control and noncontrol PRO and uses small *pro* in certain, highly restricted cases. Making use of the theoretical framework developed in Rizzi (1986), an explanation for the behavior of P *pro* and PRO with arbitrary reference will be given. Conclusions are summarized in section 4.

Data for this article has been supplied by speakers of the Aruban (A) and Curaçaoan (C) dialects of Papiamentu.¹ The A and C informants differed systematically in their opinions. Where relevant, these differences will be mentioned, and both the A and C forms of the examples will be given. Wherever there is no difference, an A orthography will be used.

2. (Non)finiteness of *pa* Complement Clauses

The literature on P, e.g., Bickerton 1984, Birmingham 1970, Dijkhoff 1983, Fodale 1984, and Washabaugh 1975, has usually assumed that *pa* introduces nonfinite clauses. This conclusion was based on the assumption that *pa*-complement clauses do not allow any markers of Tense, Mood, or Aspect (TMA). A problem with this position is that *pa* clauses often have lexical subjects. The subject position of an infinitival clause is not governed and cannot receive nominative case; hence the subject cannot be lexical. Most of the above-mentioned authors do not concern themselves with the question: how can the subject of a *pa* clause, which is assumed to be infinitival, still be lexical, thereby violating the case filter?

Case filter: *NP if NP has phonetic content and has no Case (Chomsky 1982:49).

Dijkhoff & Mufwene (1986, 1988) have provided evidence for the use of perfective aspect in *pa* clauses, thus demonstrating that the assumption that no TMA markers can appear is not correct. They argue for the position

that *P* has no finite/nonfinite distinction, hence no infinitive. Whether the subject of the complement clause is an overt subject or an [ec] subject makes no structural difference to them.

Let us consider the following *P* sentences:

- (1) a. Cheli_i a pidi Feli_j pa e_{j/k} bai playa kuné_i
 Estel_i PF ask Felix_j COMP he_j/(s)he_k go town with-she_i
 'Estel has asked Felix if he/(s)he would go into town with her.'
- b. Cheli_i a pidi Feli_j pa [ec]_j bai playa kuné_i
 Estel_i PF ask Felix_j COMP go town with-she_i
 'Estel has asked Felix to go to town with her.'

e in the examples is the third person singular (note that *P* has no overt case distinctions, hence the pronominal system shows no case agreement), [ec] is an empty category, and coreferentiality is indicated by subscript *i, j, k*. A first conclusion that can be drawn from these examples is that there is a different interpretation for *pa* clauses containing a lexical subject, as compared to those that contain an [ec] subject. In (1a) the lexical subject can be coreferential with the object of the matrix clause, or acquire its reference from outside (given the appropriate context), whereas the [ec] subject in (1b) is obligatorily coindexed with a constituent in the matrix clause.

What are the analyses that a GB framework (as developed in Chomsky 1982) provides here? If the subject of the *pa* clause is an overt NP, as in (1a), we could be dealing either with case marking by *pa*, or with a finite complement clause introduced by *pa*. Under the first option, *pa* is a prepositional complementizer, the equivalent of English *for*, introducing an infinitival clause in which the subject position is case marked by *pa*, thus licensing the overt NP subject. If the second option is true, *pa* is a genuine complementizer, and *e* is the subject of the embedded clause, receiving nominative case from governing INFL.

If the subject of the *pa* clause is an [ec], as in (1b), we could be dealing either with the PRO subject of an infinitival complement clause (PRO being the [ec] which appears in non-case-marked argument positions), or with the *pro* subject of a finite complement clause (*pro* being the [ec] that appears in case-marked positions, i.e., in positions where a lexical NP could appear). The finite/nonfinite distinction is thus invoked to account for the distribution of empty subjects and overt subjects, as well as the distribution of tense. We will come back to these analyses below.

Let us now turn to the position taken by Dijkhoff & Mufwene, that there is no finite/nonfinite distinction in Papiamentu. They base their argument on the assumption that “in the languages where it has been attested, the infinitive is a morphologically-based distinction, and that, cross-linguistically, there is no constant syntactic or semantic characteristic that can be associated with it” (1988). They claim that in sentences like the one below (1988:12), the subject can be null or lexical, the implication being that there is no real difference between *pa* clauses with PRO subject or overt NP subject²:

- (2) Maria_i ta elogiá Pedro_j pa e_i/PRO_i haña informashon
 a. Mary_i is praising Peter_j so that she_i may get information
 b. Mary_i is praising Peter_j in order PRO_i to get information

They state that the PRO/*pro* distinction does not apply to P; hence the [ec] in (2) could be either one. However, the option for the interpretation of the [ec] subject in the *pa*-complement clause as *pro* is what we would expect, as *pro* typically occurs wherever an NP subject can occur. As shown in example (1), there is a different interpretation for [ec] subjects and NP subjects in *pa*-complement clauses. In example (2) the same applies: If the subject in (2) is the pronoun *e* (third person singular), it can be coreferential with the subject of the matrix sentence, or with the object of the matrix sentence, or even with a third party, given the appropriate context:

- *(2) Maria_i ta elogiá Pedro_j pa e_{i/j/k} haña informashon
 ‘Mary_i is praising Peter_j so that she/he (s)he may get information.’

This does not hold if the subject of the embedded sentence is an [ec] subject: In that case we are looking at a subject-control structure. This difference in interpretation makes a *pro* analysis less plausible, and provides counterevidence for the assumption that there is no difference between NP subject and [ec] subject. From Dijkhoff & Mufwene’s assumption that the finite/nonfinite distinction as well as the PRO/*pro* distinction are irrelevant to P, it follows that overt and covert subjects are equal, and we can no longer explain the control differences found for the overt pronoun versus the covert pronoun.

Dijkhoff & Mufwene (1988) state that tense/aspect markers other than perfective *a* (with counterfactual interpretation) are not permitted in *pa* clauses. In the framework of this article, their analysis would make *pa* clauses infinitival, with no possibility for lexical subjects to appear. This

brings us to the option outlined earlier, that *pa* is a prepositional complementizer, with the pronoun *e* being the subject of an infinitive, and receiving case from prepositional *pa*: But is it true that the past tense marker cannot appear there? How do we ascertain the finiteness of *pa* clauses? Finiteness is indicated by the possibility of overt realization of tense. The P past tense marker is *tawata/tabata* (past imperfective). By definition, finite clauses are those that allow the tense marker to occur; when the *pa* clause is nonfinite, this option is not available.

Testing these assumptions is not easy for several reasons: the use of TMA markers in subordinate clauses is highly restricted in P, and anything other than Ø realization of TMA in *pa* complement clauses seems to represent the marked option, especially so for the A informants. Furthermore, the imperfective reading of *tawata/tabata* must fit into the propositional content of the clause, e.g., the tense marker cannot be used in the *pa* clause in (1) for this reason.

Examples (3a-b), below, are a (not entirely) successful attempt at constructing sentences that meet all the requirements; note that the A informants did not accept tense marking in the subordinate clauses at all:

- (3) a. C. mi tata_i sèmper tabata lucha p'e_i tabata
 My father always PAST struggle COMP-he PAST
 bai dilanti
 go forward
 'My father was always struggling to make progress in life.'
- b. C. *mi tata_i sèmper tabata lucha pa [ec_i] tabata
 my father always PAST struggle COMP PAST
 bai dilanti
 go forward

For the C informants, depending on TMA marking in the matrix clause and on the type of verb used, the overt tense marker can occur in the *pa* clause, provided the subject has a phonological matrix as in (3a). In example (7), below, we will see that for the A informants the use of the tense marker in a *pa* clause with overt NP subject will also be allowed; however, more restrictions are imposed.

From this we may conclude that *pa* clauses with NP subject are finite. The option of case marking by prepositional *pa* followed by an infinitival clause is hereby excluded.

As might be suspected from the coreferentiality test in (2), the [ec] subject in these *pa* clauses cannot be *pro*: *pro* would not exclude overt realization of tense where NP allows it. Wherever a tense marker is allowed to occur, PRO is excluded. This leads us to the conclusion that in the (b) examples above, *pa* introduces an infinitival clause with a controlled PRO subject. The possibility of using TMA particles in these clauses is restricted to the perfective aspect marker *a*.

It is important to note that there are further differences between the A and C speakers with respect to their judgments on the acceptability of empty subjects in *pa* complement clauses. The Arubans prefer lexical NP as subject of the *pa* clause in the examples above, even if the subject of the embedded clause is coreferential with that of the matrix clause, in which case the Curaçoans prefer PRO subject. Accordingly, PRO subjects seem to have a more marginal status for the A informants.

So far we have focused on control PRO, the content of which is determined by coreferentiality with an NP in the matrix clause. In the following section we will be discussing other empty categories in Papiamentu that can be characterized as [+ pronominal], namely, PRO with arbitrary reference (PRO_{arb}) and small *pro* (cf. Chomsky 1982 for a typology of empty categories).

3. Control PRO, Noncontrol PRO, Small pro

The sentences given below under (4) are taken from Dijkhoff & Mufwene (1988) and are representative of the C dialect of P. The clause in (4a), introduced by *pa*, has an overt pronominal subject, whereas the one in (4b) has an [ec] subject:

- (4) a. mi ke pa e piki koko
 I want COMP he pick coconut
 'I want him to pick coconuts.'
- b. mi ke pa PRO_{arb} piki koko
 I want COMP pick coconut
 'I want coconuts to be gathered.'

Even more so than in the examples (1–3) above, it is obvious that NP subject and PRO subject cannot be the same here. The arbitrary reading of PRO is obligatory, and coreferentiality with the subject of the matrix clause is excluded, as shown, for example, in (5):

- (5) a. *mi_i ke pa PRO_i piki koko
 b. mi ke pa ami piki koko
 I want COMP I_{emph} pick coconut
 'I want for me (and no one else) to pick coconuts.'

With reference to PRO_{arb} it is again important to distinguish clearly between A and C speakers of P. The example under (4a) is all right for both the C and A informants. The one under (4b), however, which is perfectly acceptable for the Curaçoans, is ungrammatical or highly marked for the Arubans. Their equivalent would be a passive construction (which is of course also a possible option in the Curaçaoan dialect)³:

- (6) A. mi kjèr pa koko wordo bahá
 I want COMP coconut PASS picked
 'I want coconuts to be picked.'

As we saw in the previous section, control PRO in infinitival *pa* clauses seems to represent a marked option for the A informants. PRO_{arb} is even more marginal in their dialect. Note that the complement clause can be tensed in (4a) (which has a lexical subject in the *pa* clause). This is an example where Aruban Papiamentu also allows tense marking in the *pa* clause, as exemplified under (7).⁴ In (4b), however, where the embedded clause has a PRO subject, tense marking is not allowed for either dialect, as shown under (8).

- (7) A. mi tawata kjèr pa bo tawata baha koko
 C. mi tabata ke pa bo tabata piki koko
 I PAST want COMP you PAST pick coconut
 'I wanted you to have been gathering coconuts.'
- (8) C. *mi tabata ke pa PRO_{arb} tabata piki koko

This again confirms the earlier established finite/nonfinite distinction for *pa* clauses, and provides further evidence against Dijkhoff & Mufwene's contention that the distinction PRO/*pro* is irrelevant for P.

Even though P is not a null subject language (we have seen in the foregoing that *pro* cannot simply occur wherever an NP subject can), it sometimes allows for small *pro* in the subject position of a tensed sentence. We can distinguish among the following usages of *pro* in P:

- a. *Nonargumental*: Where English has expletive *it* to serve as the dummy subject of raising verbs such as *seem*, P has an obligatory [ec] subject, as demonstrated under (9).

- (9) *pro* (tawata) parse ku Maria ta(wata) malo
 (PAST) seem COMP Mary be(PAST) ill
 'It seems/seemed that Mary is/was ill.'

Note that 'seem' is not a raising verb in P; hence, "Mary seems to be ill" is ungrammatical in P. Other instances of nonargumental usage of *pro* are:

- (10) *pro* tin /tawatin hopi hende
 have/PAST-have many person
 'There are/were many people.'
- (11) *pro* (tawata) parse ku *pro* (taw)tin hopi hende
 (PAST) seem that (PAST)have many person
 'It seems(seemed) that there are(were) a lot of people.'

b. *Quasi argumental*: Weather verbs such as *jobe* 'rain' have an implicit subject (hence the term "quasi argumental"), in this case *awa* 'water', which can be, but does not need to be, expressed overtly. The A informants prefer the overt version, the Curaçaoans the covert one:

- (12) A. ?*pro* tawata jobe vs.: awa tawata jobe
 C. *pro* tabata jobe
 PAST rain
 'It was raining.'
- (13) A. ?ajera mi tawata kjèr pa *pro* tawata jobe
 C. ajera mi tabata ke pa *pro* tabata jobe
 yesterday I PAST want COMP PAST rain
 'Yesterday I wanted it to rain.'

Note that *pa* in (13) introduces a tensed clause with *pro* subject. This demonstrates again that a finite/nonfinite distinction needs to be made to account for the different distribution of PRO/*pro*, as PRO (controlled or arbitrary) cannot occur as subject of a weather verb:

- (13) a. **mi*_i kjèr pa PRO_{i/arb} jobe
 I want COMP rain

c. *Arbitrary*: In (14), below, *pro* is used with arbitrary reference in the argument position of a verb which usually takes a lexical NP subject. The example is acceptable to the Aruban informants only as an announcement on a flowershop, preferably with some extras like *aki* 'here' and *bunita* 'pretty', plus a big exclamation mark, as in (14b). It is not acceptable to them as a description of ongoing activity. (15) is unacceptable because we

need to specify who is playing (as for example *nan* 'they').

- (14) a. *pro* ta bende flor
 PRES sell flower
 C. 'Flowers are being sold.'
 A. 'flowers for sale'
- b. aki ta bende flor bunita!
 here PRES sell flower pretty
 C. 'Pretty flowers are being sold here.'
 A./C. 'pretty flowers for sale here'
- (15) A. **pro* ta(wata) toka bon musika
 C. *pro* ta(bata) toka bon musika
 PRES/PAST play good music
 'Good music is/was being played.'

The missing subject in languages such as Portuguese and Spanish (the lexifier languages of P), which belong to the class of pro-drop languages, is usually sufficiently compensated for by a "rich enough" inflectional system, so that the inflection determines the grammatical features of the "missing subject" (Chomsky 1982:86). The content of the empty subject must be determined by agreement (AGR), which is described as "a set of specifications for the features person, gender, number, and (in pro-drop languages) Case" (Chomsky 1982:87). In these languages, *pro* functions as an empty pronoun.

P, on the other hand, has very little verb morphology, so that the person and number features of the empty pronoun cannot be recovered from this source. The fact that we can still have a missing subject in the examples given here is due precisely to the "emptiness" of these positions in terms of their semantic content. This is especially clear in the dialect of the A informants, who seem to impose even more restrictions on the use of *pro* than do the informants from Curaçao. The nonargumental usage of small *pro* is acceptable (examples [9–11]), but its quasi argumental (and especially its arbitrary) usage pose some problems. The Curaçaoans allow for a wider definition of semantic emptiness, including arbitrary usage, but they still exclude referential usage.

If we want to explain the behavior of P small *pro*, the assumptions quoted above from Chomsky (1982) are insufficient. They do not allow for any differentiation in the features that determine the content of *pro*, and therefore cannot account for the fact that: (1) P *pro* does not function as an

empty pronoun as it does in null subject languages, (2) speakers of different dialects of P differ in their judgments on the acceptability of *pro*, and (3) there seems to be a hierarchy in these judgments related to the levels of semantic "emptiness" of *pro*.

Rizzi (1986), in describing Italian small *pro* and comparing the Italian case with that of some other languages, suggests the following solution to this problem: The content of *pro* must be recoverable, but languages may differ in the features which play a role in this procedure. In other words, a particular language may not use all of the features mentioned above (person, gender, number, Case). He also suggests that nonarguments, quasi arguments, and referential arguments are hierarchically ordered with respect to the required "richness" of their content, expressed in terms of their features (Rizzi 1986:543). In other words, nonarguments, quasi arguments, and referential arguments do not all require the same features to be specified. The absence of morphological person-marking on the verb follows if we assume that P AGR is weak, which means that AGR is not specified for person and number. Note that we cannot assume that AGR is empty (as Dijkhoff & Mufwene 1988 do), as this leaves us with no means to account for the fact that *pro* in P is allowed to occur only in the cases exemplified above. If AGR were empty, we would either expect no licensing of *pro* at all, or licensing "across the board."

Let us consider the features that play a role in the way P *pro* is used: the nonargumental usage of *pro*, which is the semantically emptiest one, requires no features at all. The quasi argumental usage of *pro* (i.e., the weather verb cases) needs no specification for number: the implicit subject in examples like (12) and (13) is a noncount noun (i.e., *awa* 'water'). The arbitrary usage of *pro* has [+human] gender specification, but it has no specification for person, nor does it have one for number — nothing is implied about singular or plural subjects, as in example (14–15), where one or more persons can be involved in selling flowers and playing music.

We suggest the following: AGR is extremely weak in P, and can only be specified for the feature gender as [+human] or [–human], but it is not necessarily specified for this feature. If no content of AGR is recovered, in other words, if AGR has no specification for gender, *pro* is nonargumental. If gender is specified as [–human], then *pro* is quasi argumental. If there is a specification for gender as [+human], then we are dealing with arbitrary *pro*.

In Rizzi's model, each individual language also has to decide what

belongs to the set of licensing heads for small *pro*. As P allows only for *pro* in the subject position of tensed clauses, only strong INFL (i.e., [+finite] INFL) is a licensing head.

In order to account for the different judgments of A speakers on the acceptability of small *pro*, as compared to Curaçaoan speakers, we will propose that the ability to specify the AGR feature gender is weak or does not exist at all for the Aruban speakers; therefore the arbitrary interpretation of *pro* cannot unambiguously be distinguished from its quasi argumental usage, or even from its nonargumental usage.

This also offers an explanation for the restrictions on the use of PRO with arbitrary interpretation for the A informants: *arb* has to be specified as [+human]. If Aruban Papiamentu lacks the ability to specify the gender feature, then the [+human] specification of PRO has to be recovered through the option of control. In addition, we may have to assume that the specifications for the recovery procedure suggested by Rizzi for *pro* also hold for PRO in P.

4. *Conclusions*

It has been shown that there are syntactic and semantic tests which provide proof for a distinction between finite and nonfinite clauses introduced by the complementizer *pa*. In addition, we have seen that the finite/nonfinite distinction accounts for the distribution of empty and overt subjects of *pa* clauses. Finite complement clauses are those in which the subject position is lexical and the overt tense marker *tawa/tabata* may be used, although its usage is severely restricted, as it is dependent on TMA marking in the matrix clause, on the type of verbs used, and on restrictions on TMA marking which may be specific to the dialect of the speaker. Needless to say, these phenomena call for further research. Infinitival complement clauses are those where the subject position is filled with PRO and the overt tense marker may never appear.

P allows for small *pro* as the subject of a tensed sentence, i.e., it is governed by strong INFL. The analysis presented here follows Rizzi (1986) in assuming that governing INFL licenses *pro*, and that *pro* therefore has the grammatical specifications of the features of agreement. As AGR is weak in P, *pro* cannot receive person or number specification, and it must be restricted to nonreferential usage. We can distinguish among arbitrary, quasi argumental, and nonargumental use. The content of P *pro* is almost

empty. In the recovery procedure only the feature gender (i.e., [+human] or [-human]) plays a role, and it may be that in Aruban Papiamentu not even the AGR feature, gender, can be specified. Thus the different interpretations of *pro* cannot be unambiguously distinguished from each other, and from this it follows that quasi argumental and arbitrary use of *pro* is marginal in the A dialect.

P has control PRO and noncontrol PRO (PRO_{arb}). Both have marginal status in Aruban Papiamentu. For the recovery of the content of PRO with arbitrary interpretation, [+human] gender again needs to be specified. It may be that Aruban Papiamentu lacks this possibility as well.

Since there are syntactic differences between Aruban and Curaçaoan Papiamentu, it is important that authors specify the dialect with which they are concerned. In this paper, major differences were noted in restrictions on TMA marking in *pa* complement clauses, the status of control and non-control PRO, and the status of small *pro*.

NOTES

This article is a revised version of a paper presented at the Creole Workshop, held March 23–24, 1987, at the University of Amsterdam. I have benefited from suggestions and comments made by many people, in particular Peter Bakker, Hans den Besten, Sabine Iatridou, Carla Luyks, and Pieter Muysken.

1) Many of my Curaçaoan and Aruban friends have supplied data for this article and have helped with the interpretation, especially Franklin Benito (C.), Haime Jones (A.), Homen Jeung (C.), and Ronnie Kock (A.).

Although Bonairan Papiamentu is generally believed to be very similar to the Curaçaoan dialect, the conclusions drawn in this paper are not necessarily valid for Bonairan.

2) In their 1988 paper, Dijkhoff & Mufwene contrast *pa* clauses with *ku* clauses, and they compare their data from P with that from other languages. I will not do so in this article, and instead compare the behavior of empty subjects in *pa* clauses with empty subjects in other positions in P.

3) The fact that P has a passive construction makes an interpretation of (4b) as a passive unlikely, in my opinion. Besides, **piki koko* for 'coconuts have been picked' is ungrammatical.

4) More work remains to be done on the restrictions on tense marking in subordinate clauses in Papiamentu. Besides the factors mentioned here, the categorial status of the matrix verb seems to play a role.

REFERENCES

- Bickerton, Derek. 1984. The language bioprogram hypothesis. *The Behavioral and Brain Sciences* 7.173–221.
- Birmingham, J. C. 1970. The Papiamentu language of Curaçao. Charlottesville: University of Virginia dissertation.
- Chomsky, Noam. 1982. Some concepts and consequences of the theory of Government and Binding. Cambridge, MA: MIT Press.
- Dijkhoff, Marta B. 1983. Movement rules and the resumptive pronoun strategy in Papiamentu. University of Groningen master's thesis.
- , and Salikoko Mufwene. 1986. Notes on the so-called “infinitive” in creoles. Sixth Biennial Meeting of the Society for Caribbean Linguistics, St. Augustine, Trinidad. ms.
- , and ———. 1988. On the so-called infinitive in Atlantic Creoles. ms.
- Fodale, Peter. 1984. Non-finite sentential complementation in Papiamentu. *Papers in Linguistics* (University of Michigan) 5:1.1–15.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17.501–59.
- Washabaugh, William. 1975. On the development of complementizers in creolization. *Working Papers in Language Universals* 17.109–40.